

Progress Report and Planning

Task 3 – Web Community Sensing

Eduarda Mendes Rodrigues

REACTION Workshop, FCUL

January 6th, 2011

Web Community Sensing

- Modeling the credibility and authority of news sources and **opinion makers in social networks**
- Identifying **influential individuals and experts** on a given news topic
- Monitoring the **community reaction to news** stories and the polarity of opinions

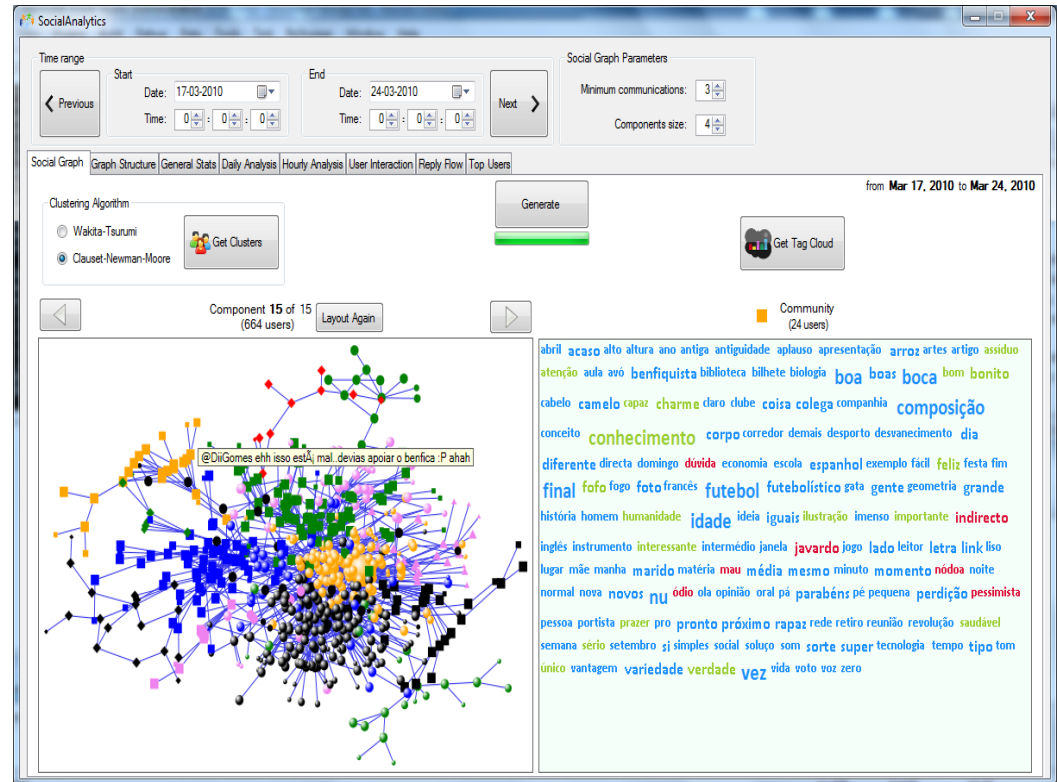
Progress Report

- Data collection, pre-processing and exploratory analysis
 - news comments, blogs, micro-blogs
 - pre-processing of Twitter messages (e.g., tokenization)
 - analysis of the Twitter social network (SocialAnalytics tool)
- Influence detection algorithms (MSc project)
 - implementation based on prior research
 - prototype to demonstrate ranking of influential micro-bloggers in relation to particular news themes
- Reference corpora for evaluation (MSc project)
 - development of an annotated social graph
 - web interface for management of the evaluation resources
- Hunting for HR

SocialAnalytics Tool

Exploratory analysis of social network properties

- Community detection by topic
- User centrality analysis
- Social network structure visualization



Human Resources

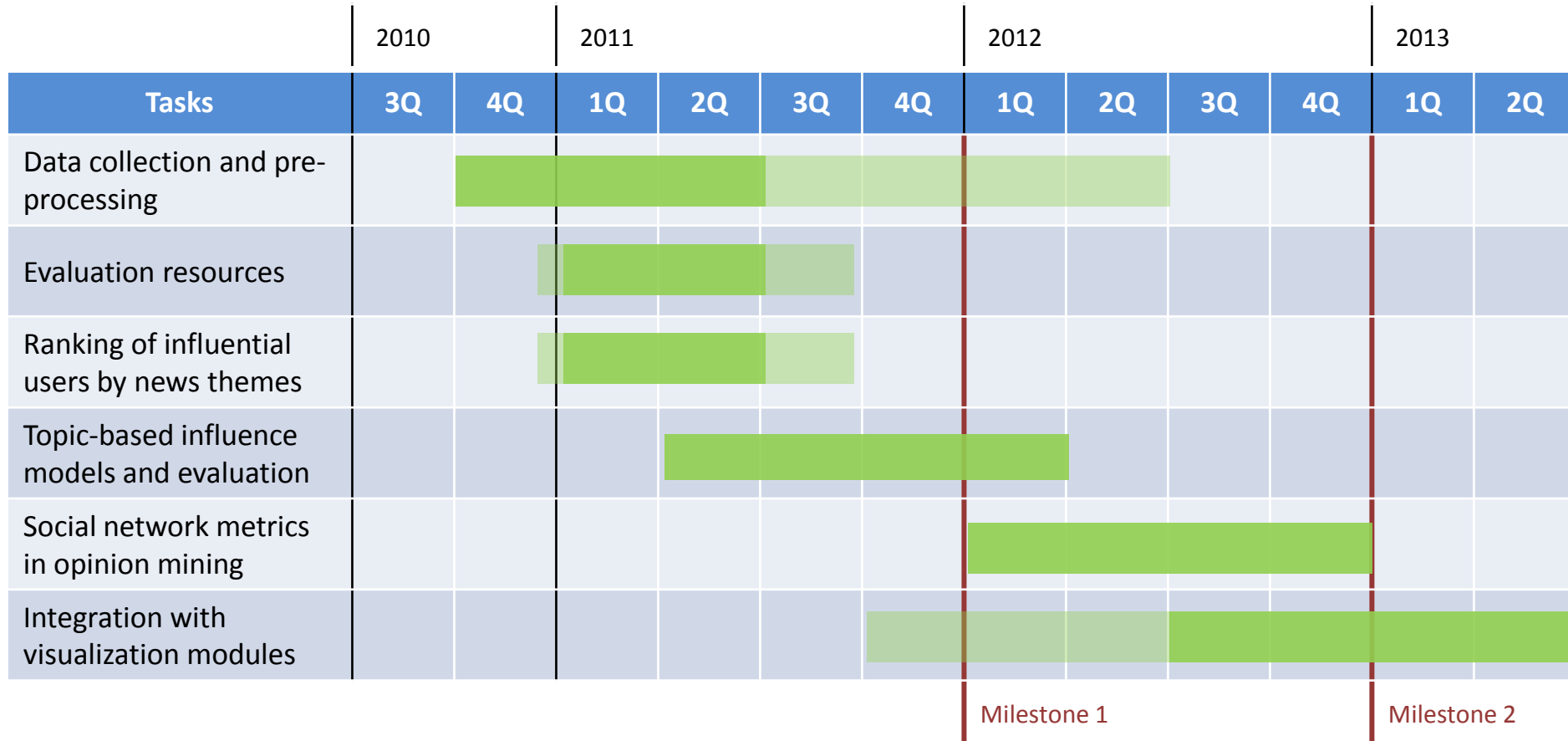
- 1 PhD Student
 - Gustavo Laboreiro (Twitter data collection and pre-processing)
- 2 MSc Students
 - José Martins (implementation of influence ranking methods)
 - Eduardo Oliveira (development of evaluation resources)
- BI9: MSc Graduate/PhD Candidate^{*}
 - Currently hiring (completed first round of interviews)
 - Expected start date: mid-Feb / Mar

^{*} FCT approved the following changes to the BI grants allocated to task 3:
3 BI BSc (12 months) → 2 BI MSc (12 months) + 2 BIC (3 months)

BI9 Candidates

- Matko Bošnjak
 - Prior experience in text mining (named entity extraction) and recommender systems
 - Strong background in machine learning
 - Versatile programmer
 - Potential PhD candidate
- Tanmoy Mukherjee
 - Prior experience in influence modeling in social networks
 - Strong background in machine learning
 - Already enrolled in a PhD program

Task 3 Timeline



Milestone 1: Initial Specification of Computational Newsroom

Identification of the requirements for the content analysis, web analysis and query and visualization tools to be later installed in the Computational Newsroom. Initial prototypes of key software components will be demonstrable.

Milestone 2: Demonstrable Computational Newsroom

Demonstration of the operation of the Computational Newsroom showcasing the results of the conducted research and providing evidence of its advantages in comparison with conventional journalism.